**DNA Analysis**

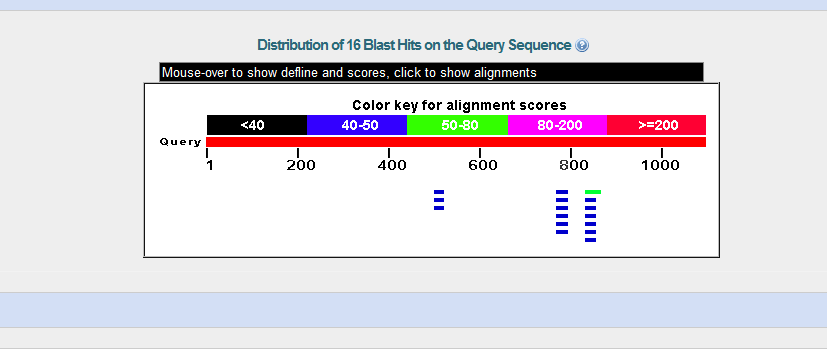
Result of Sequencing:

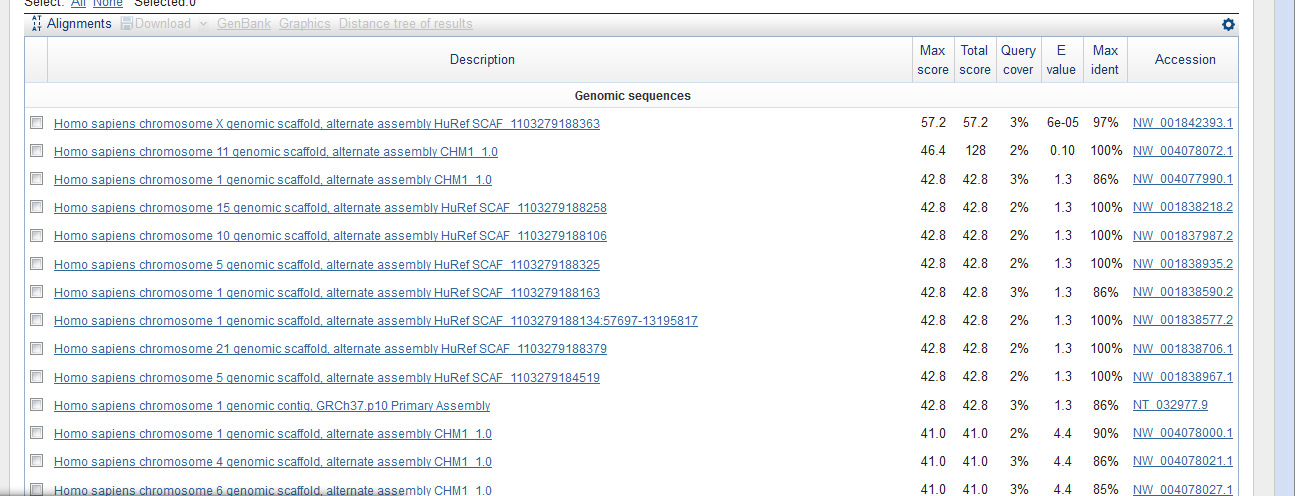
NNNNNNNNNNNTATNGNATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGTGATTTTGATTGATTGAAGGAGAAATATCATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTGTCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAGCAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATCTCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCACAACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTCGGTCATCACCATCACCATCACTAAAATCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAANAGGCCCGCACCGATCGCCCTTCCCACAGTTGCGCAGCCTGATGGCNATGGACGCNNCCTGNANCGNCGCATTAGCNCGGCNGGNNNGNGNNCNCGCAGCGNGACNCTACANTTNCNGCNNCNANCNNNNNCNTTCNNTTNTNCNNNTNNNNNNNNNNNNNNNTTNCCGNNAGNTNNAANNGGGGNNNNNNNNNATNNGNTNNNNNNNNNNCCAAANTGNNNGNNNNGNNNNNNNNGNNNNNNNNNNNNNNNNNNNNNNNGANCNNNNNNNNNNNNANNNNNNNNNNNCNNNNNAANNNNNCCNNTNNN

Sequence without extraneous N’s:

TATNGNATACTCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCACTAGTGATTTTGATTGATTGAAGGAGAAATATCATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTGTCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAGCAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATCTCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCACAACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTCGGTCATCACCATCACCATCACTAAAATCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAANAGGCCCGCACCGATCGCCCTTCCCACAGTTGCGCAGCCTGATGGCNATGGACGCNNCCTGNANCGNCGCATTAGCNCGGCNGGNNNGNGNNCNCGCAGCGNGACNCTACANTTNCNGCNNCNANC

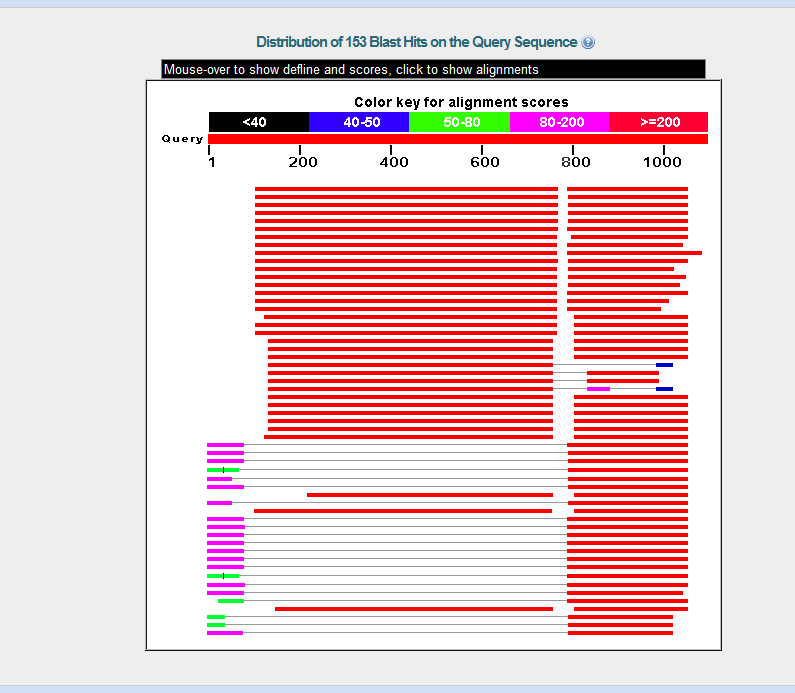
Human Genome BLAST Results:

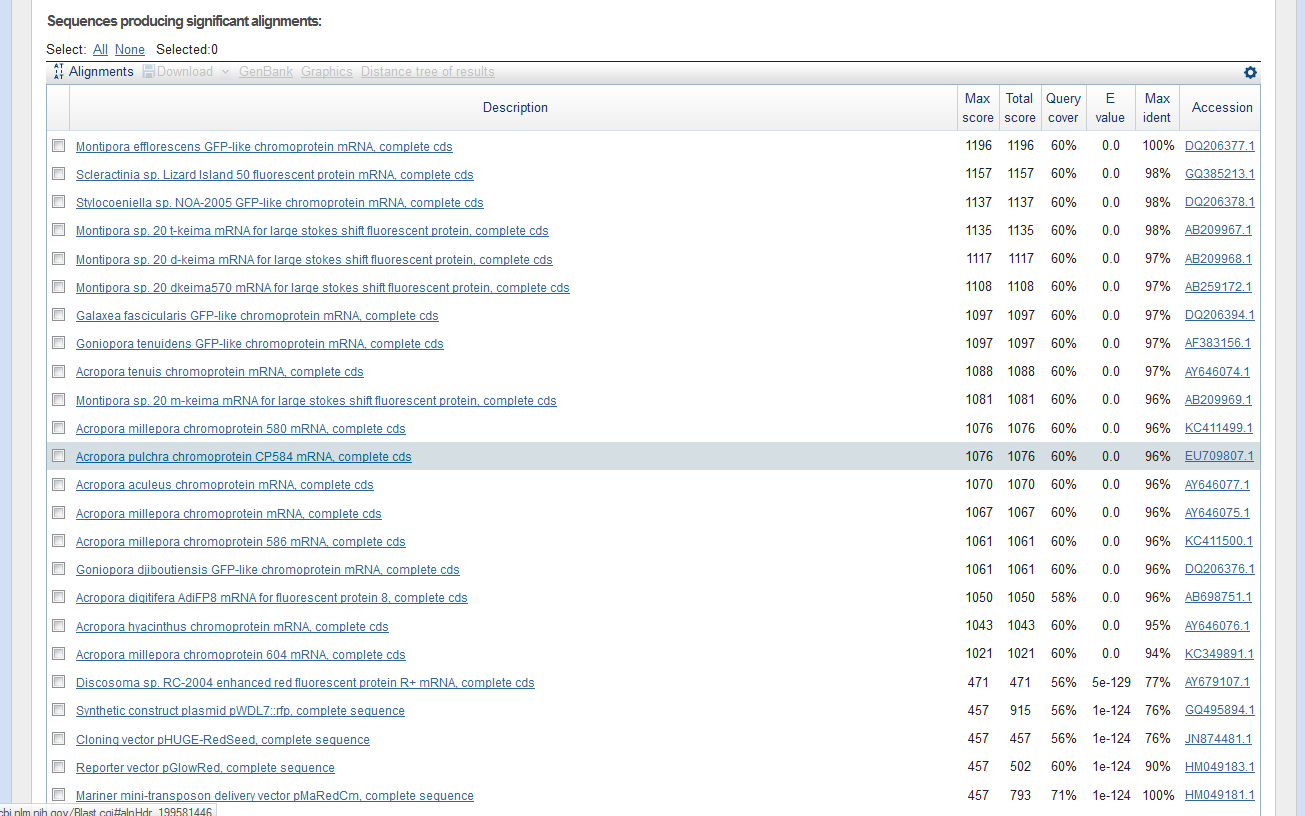




The above are the results from the BLAST when comparing to Human genome. The results show that there is not a match for the sequence provided. This is due to the fact that humans don’t have the pGBR22 gene, which would make humans fluoresce.

Result from BLAST with Other Genomes:





These are the results from the BLAST when comparing to all organisms. It shows that the provided sequence is a 99% match to monopora efforescens GFP-like chromoprotein mRNA complete cds.

Comparison to Known pGBR22 Gene:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1258 bits(681) | 0.0 | 683/684(99%) | 0/684(0%) | Plus/Plus |

Query 1 ATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGA 60

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 105 ATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGA 164

Query 61 CACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTA 120

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 165 CACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTA 224

Query 121 AAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTG 180

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 225 AAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTG 284

Query 181 TCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAG 240

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 285 TCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAG 344

Query 241 CAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTG 300

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 345 CAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTG 404

Query 301 TGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATC 360

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 405 TGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATC 464

Query 361 TCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAA 420

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 465 TCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAA 524

Query 421 CCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCT 480

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 525 CCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCT 584

Query 481 CTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAGTTCAAATCTACTTACAAGGCAAAG 540

||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||

Sbjct 585 CTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAG 644

Query 541 AAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCAC 600

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 645 AAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCAC 704

Query 601 AACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTC 660

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Sbjct 705 AACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTC 764

Query 661 GGTCATCACCATCACCATCACTAA 684

||||||||||||||||||||||||

Sbjct 765 GGTCATCACCATCACCATCACTAA 788

There is a mismatch where a G was replaced with an A at about 515 and 619. This could be an issue, but it depends on how essential this base pair is, which is dependent on its location in the codon and if it is needed for a gene.

Translation of Sequence:

Frame 1:

YXILKLCIQRVGSSPIWSTCRRPH\*\*F\*LIEGEIS\*V\*SLNK\*PTRFICQARSMDTTLRS

KAMEKESLTRGSRR\*SSLSPRVDLCHLLGIFYHHCLNTEAYHSPSTLKTSLIM\*SSHSLR

DIHGRGS\*TLKMVQCVLSAMIPASKATVSSTMSKSLV\*TFLPMDLLCRRRHRAGNPTLSV

SLHEMEC\*\*ETTLWL\*SWKEVVTICVNSNLLTRQRSL\*GCQGITMLTANWM\*PVTTRITH

LLSSVKYPLHATLCSVITITITKIPRPWRPGACDVGPNSPYSESYYNSLAVVLQRRDWEN

PGVTQLNRLAAHPPFASWRNSEXARTDRPSHSCAA\*WXWTXPXXXH\*XGXXXXXQRDXTX

XAXX

Frame 2:

XXYSSYASNALGALPYGRPAGGRTSDFD\*LKEKYHECDR\*TNDLQGLYVRHGQWTLL\*GR

RRWKRKALRGGADGKAHCHQGWTSAICLGYFITTVSIRKHTIHQVP\*RHP\*LCKAVIP\*G

IYMGEDHEL\*RWCSVYCQQ\*FQHPRQLFHLQCQNLWCELSSQWTCYAEEDTGLGTQH\*AS

LCTRWNADRKQLYGSEVGRRWLLFV\*IQIYLQGKEACEDARVSLC\*PQTGCNQSQQGLHI

C\*AV\*NIHCTPLFARSSPSPSLKSRGHGGREHATSGPIRPIVSRITIHWPSFYNVVTGKT

LALPNLIALQHIPLSPAGVIAXRPAPIALPTVAQPDGXGRXLXRRISXAGXXXRSXTLXX

XXX

Frame 3:

XXTQAMHPTRWELSHMVDLQAAALVILID\*RRNIMSVIAKQMTYKVYMSGTVNGHYFEVE

GDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPLSQYGSIPFTKYPEDIPDYVKQSFPEG

YTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKISGVNFPPNGPVMQKKTQGWEPNTERL

FARDGMLIGNNFMALKLEGGGYYLCEFKSTYKAKKPVRMPGYHYVDRKLDVTSHNKDYTF

VEQCEISIARHSLLGHHHHHH\*NPAAMAAGSMRRRAQFAL\*\*VVLQFTGRRFTTS\*LGKP

WRYPT\*SPCSTSPFRQLA\*\*RXGPHRSPFPQLRSLMAMDAXXXXALARXXXXXAAXXYXX

XXX

Frame 3 seems to be the only frame that would make sense since it’s the only one with the start codon and the His tag. Next, the known sequence for pGBR22 was translated (below).

Frame 1:

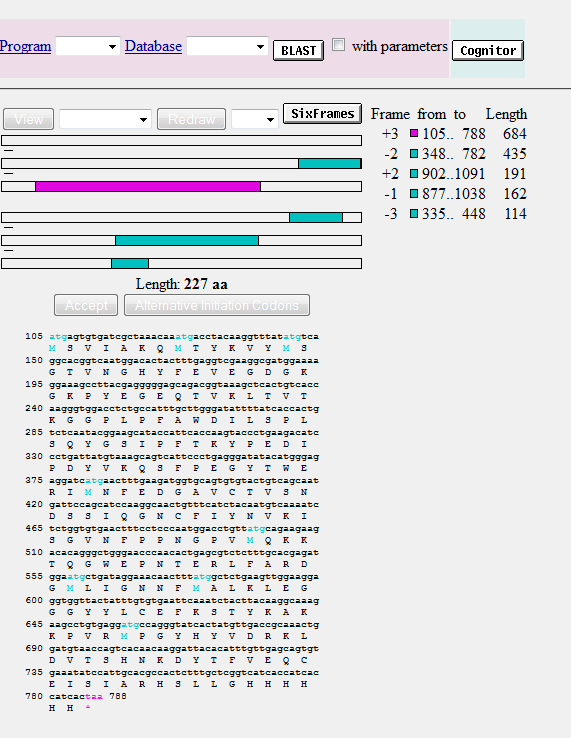
MSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPL

SQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKI

SGVNFPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGYYLCEFKSTYKAK

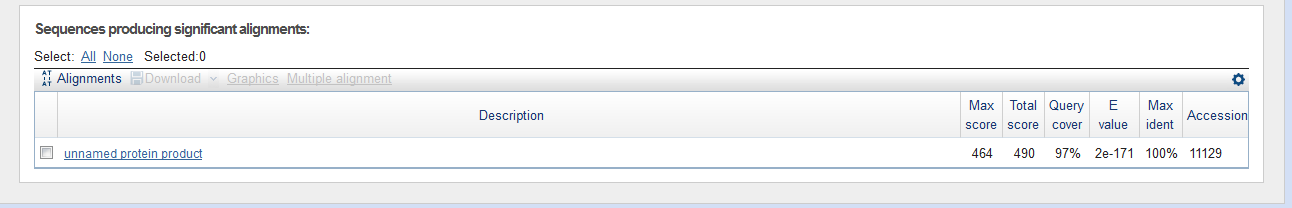
KPVRMPGYHYVDRKLDVTSHNKDYTFVEQCEISIARHSLLGHHHHHH\*

Below are the result from the ORF Finder. It is apparent that the third option is the best for this sequence. The third option is the second reading frame forward.



Protein Matching:

Next the known protein sequence of pGBR22 and the obtained sequence are compared.



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Score** | **Expect** | **Method** | **Identities** | **Positives** | **Gaps** |
| 464 bits(1194) | 2e-171 | Compositional matrix adjust. | 222/222(100%) | 222/222(100%) | 0/222(0%) |

Query 30 MSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPL 89

MSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPL

Sbjct 1 MSVIAKQMTYKVYMSGTVNGHYFEVEGDGKGKPYEGEQTVKLTVTKGGPLPFAWDILSPL 60

Query 90 SQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKI 149

SQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKI

Sbjct 61 SQYGSIPFTKYPEDIPDYVKQSFPEGYTWERIMNFEDGAVCTVSNDSSIQGNCFIYNVKI 120

Query 150 SGVNFPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGYYLCEFKSTYKAK 209

SGVNFPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGYYLCEFKSTYKAK

Sbjct 121 SGVNFPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGYYLCEFKSTYKAK 180

Query 210 KPVRMPGYHYVDRKLDVTSHNKDYTFVEQCEISIARHSLLGH 251

KPVRMPGYHYVDRKLDVTSHNKDYTFVEQCEISIARHSLLGH

Sbjct 181 KPVRMPGYHYVDRKLDVTSHNKDYTFVEQCEISIARHSLLGH 222

The results show that there is a 100% match between the two sequences. This also shows us that the mismatch in the DNA sequence of the sample, does not affect the protein and hence is not an issue.

Insertion of Vector:

Below is the DNA sequence of pGEMT, obtained from Google Drive, with the numbers removed by using Sequence Manipulation Suite.

GGGCGAATTGGGCCCGACGTCGCATGCTCCCGGCCGCCATGGCCGCGGGATTAATCACTAG

TGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTG

AGTATTCTATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTG

TGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAA

GCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCT

TTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGA

GGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTC

GTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAA

TCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT

AAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAA

AATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTT

CCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTG

TCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTC

AGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCC

GACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA

TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT

ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATC

TGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA

CAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAA

AAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAA

AACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTT

TTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGAC

AGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCC

ATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGC

CCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATA

AACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATC

CAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGC

AACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCA

TTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAA

GCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCA

CTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTT

TCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGT

TGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTG

CTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGA

TCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACC

AGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCG

ACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAG

GGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGG

GTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGATGCGGTGTGAAATACCGCACAGATG

CGTAAGGAGAAAATACCGCATCAGGAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCG

TTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCT

TATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGT

CCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGAT

GGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCA

CTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAAC

GTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTA

GCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCG

TCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGC

TATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAG

GGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTAT

A

**Reverse Complement Strand:**

TATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC

CCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAAT

AGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGG

ACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCG

CTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCA

CGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTA

GTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGC

CATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTG

GACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTAT

AAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTA

ACGCGAATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGGTATTTTCTCCTTAC

GCATCTGTGCGGTATTTCACACCGCATCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA

CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAAC

CCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTG

TCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGC

TGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGG

ATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGA

GCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGC

AACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAG

AAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGA

GTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCG

CTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGA

ATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGT

TGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACT

GGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGT

TTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGG

GGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTA

TGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAC

TGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTA

AAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGT

TTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTT

TTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTT

GTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGC

AGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTG

TAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCG

ATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGT

CGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAAC

TGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGG

ACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGG

GAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGAT

TTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTT

TACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTG

ATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAA

CGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGC

CTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGA

AAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGG

CTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTC

ACA**CAGGAAACAGCTATGAC**CATGATTACGCCAAGCTATTTAGGTGACACTATAGAATAC

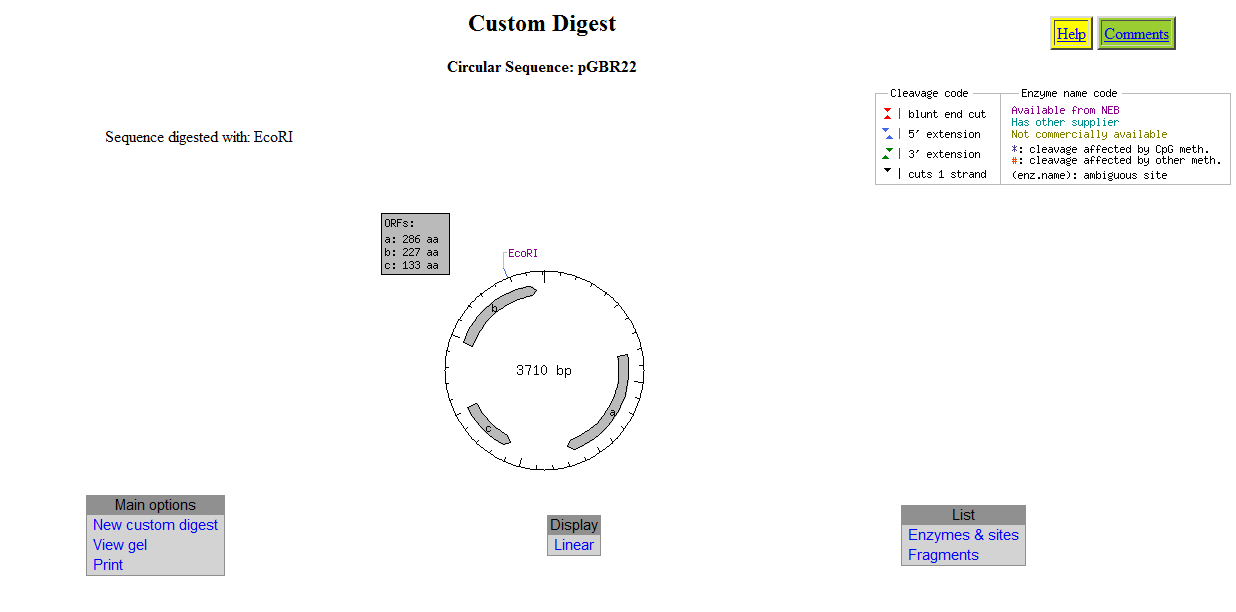
TCAAGCTATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGC

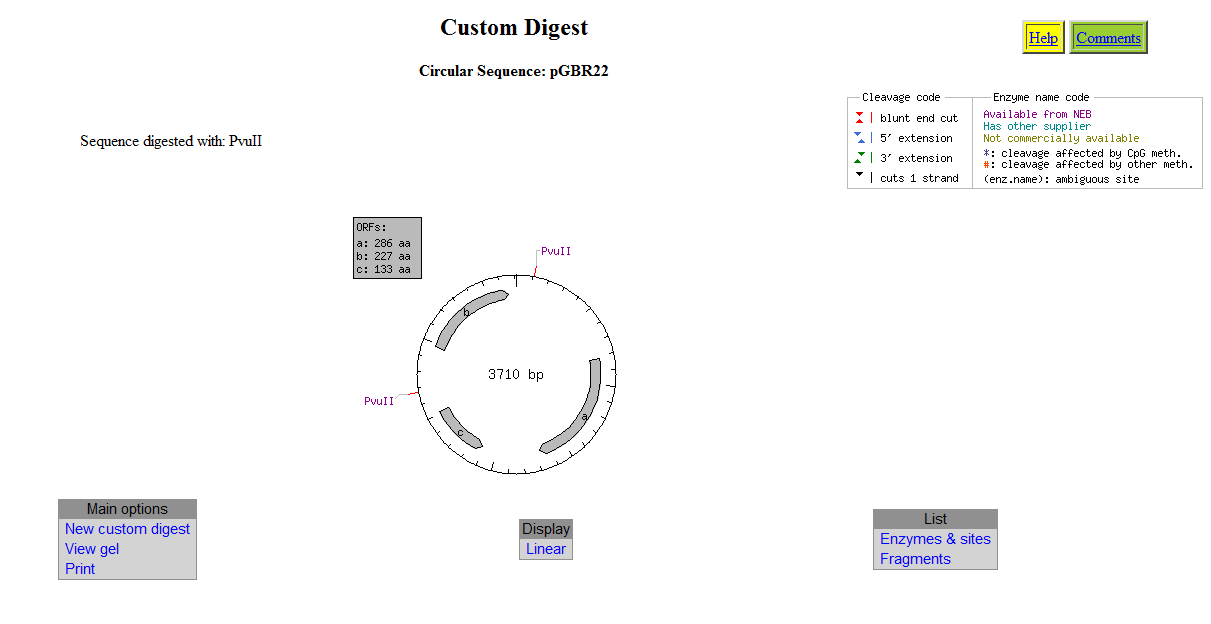
ACTAGTGATTTTGATTGATTGAAGGAGAAATATCATGAGTGTGATCGCTAAACAAATGACCTACAAGGTTTATATGTCAGGCACGGTCAATGGACACTACTTTGAGGTCGAAGGCGATGGAAAAGGAAAGCCTTACGAGGGGGAGCAGACGGTAAAGCTCACTGTCACCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACTGTCTCAATACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGATTATGTAAAGCAGTCATTCCCTGAGGGATATACATGGGAGAGGATCATGAACTTTGAAGATGGTGCAGTGTGTACTGTCAGCAATGATTCCAGCATCCAAGGCAACTGTTTCATCTACAATGTCAAAATCTCTGGTGTGAACTTTCCTCCCAATGGACCTGTTATGCAGAAGAAGACACAGGGCTGGGAACCCAACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAGTTGGAAGGAGGTGGTTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAAGAAGCCTGTGAGGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAGTCACAACAAGGATTACACATTTGTTGAGCAGTGTGAAATATCCATTGCACGCCACTCTTTGCTCGGTCATCACCATCACCATCACTAAAATCCCGCGGCCATGGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGC

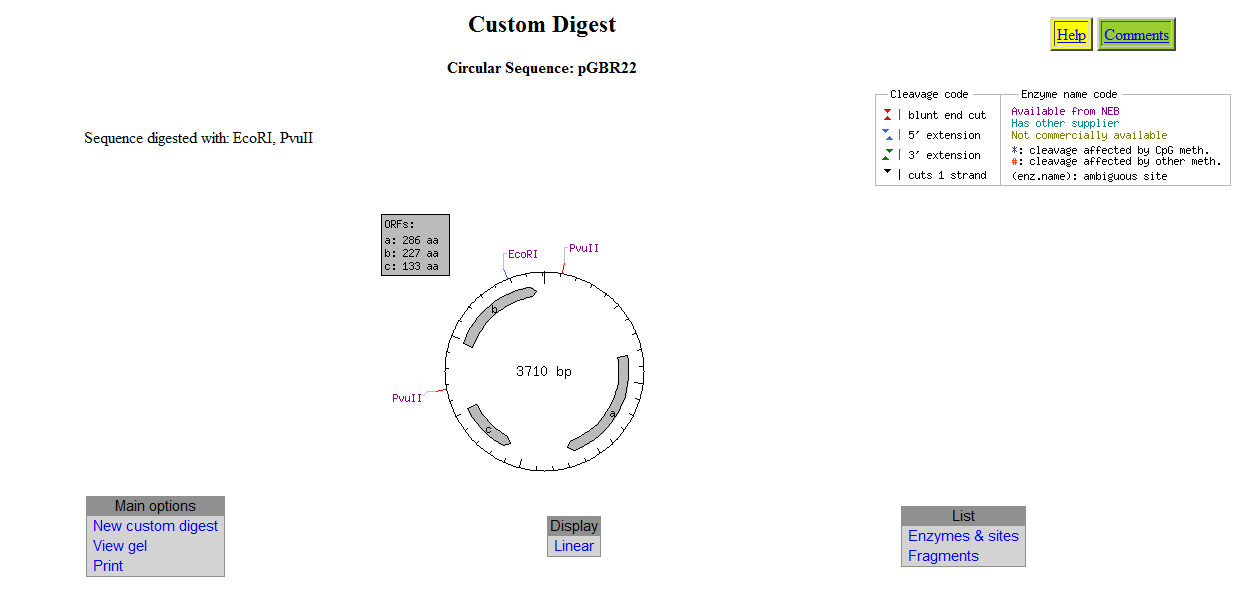
CC

The above is the reverse complament strand with the Shine-Dalgarno sequence highlighted in red, the start codon highlighted in green, and the end codon in blue. The underlined section is the inserted gene into the pGEMT sequence. The bolded underlined sequence shows where the M13 Reverse primer would bind.

Enzyme Digest:

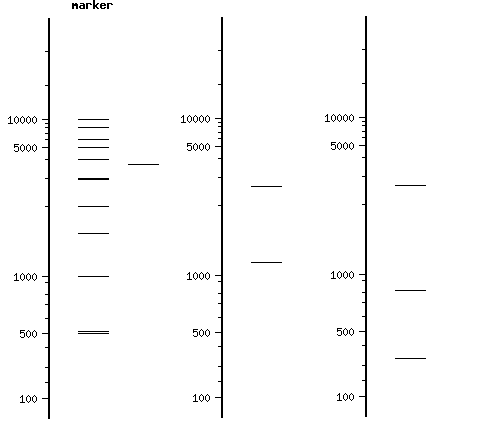






The above show the custom digest of EcoRI, PuvII and both at the same time. Virtual gels of these three were also ran (shown bellow) and used to compare the results that will be obtained in wet lab.

Virtual Gel:

 The first lane is the 1 kb DNA Ladder, the second is the EcoRI, the third is the PUVII, and the fourth has both EcoRI and PuvII.